

Maher Asaad Baker

HOW
I wrote a
million
Wikipedia
articles

Second edition

How I wrote a million Wikipedia articles

How I wrote a million Wikipedia articles

Second edition

ISBN: 978-3-7554-1660-9

München : BookRix

Publication Date: 07-01-2022

Maher Asaad Baker

Copyright © 2022 Maher Asaad Baker

All rights reserved. No part of this document may be reproduced or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission of the publisher.

DEDICATION

My mother, I still remember my first fair.

My father, I guess reading is in our blood.

Aber, Yazan, and Solav, My true love, my true stars.

Dear, Still looking for a word to describe you.

Obada, I'm still writing code.

Fadi, My true friend when I needed one.

Khalil, Best money maker.

Ghaly, Raafat and Fuad, Thanks for being there.

CONTENTS

THE PROMISE.....	2
1977	7
IN RIYADH	14
BRIGHTNESS COLORS.....	22
BACK TO SYRIA.....	26
WIKIPEDIA.....	30
FIRST WIKIPEDIA ARTICLE	36
EGYPTIAN WIKIPEDIA	39
WRITING MORE	45
HOW TO	53
WHAT TO WRITE.....	66
FINAL WORDS	73
FURTHER READING	79
DISCLAIMER	81
ABOUT THE AUTHOR	84

ACKNOWLEDGMENTS

Although in this book I will tell you how I collected and managed the data, the way I prepared the articles, and the way I automated the process to finally achieved adding 1 million Wikipedia articles, it won't be a step-by-step tutorial book, this book is a more biographical story about a period of my life, it's not even the full story, you'll see how through the struggles and pain that I lived and still living, I made something that I'm so proud of, something that maybe exists for centuries, maybe not, at least I made it, this is the story that led me to added 1 million articles to Egyptian Wikipedia.

This book was not written by an expert nor native English speaker, so don't take it as a reference to what you should do, and yes, you'll find typos and/or sentences are Grammarly wrong, keep in mind that this book is about what I did, not about what you should do and not a call for you to do it, your actions are your own responsibly.

MY STORY

THE PROMISE

In 2020 I made a promise to Raafat (which is a sysop on Egyptian Wikipedia), I promised him that I'll grow up Egyptian Wikipedia by adding 1 million articles to it,

sometimes I keep my promises, this time it was more like a challenge more than it was a promise, I challenged myself at the first place, and secondly I challenged everyone who was, and/or still against the project, saying, yes we can do it, eventually I've kept my promise, Egyptian Wikipedia is now over 1 million articles.

But before I even tell you how did I do it, let me first tell you who am I, and what cause me to do what I did, because every story should start somewhere, and every action must be driven by some purpose, however, my reason may seem different, simply, I was a programmer who was dying and decided to do something, so this is the story of me changing over years, this is a story that

goes back to 1997.

If you are reading this book to just know how I did it, you can skip these parts, but you'll miss the fun, I guess! You'll miss the important part of the telling, otherwise, just relax and get ready to know the true things behind the scenes.

**"You need something to open
up a new door, to show you
something you have seen
before but overlooked a
hundred times or more"**

Bob Dylan

1977

I've been a programmer for 24 years now, it all began in 1997 when I wrote my first software, we used to call it software then, now everything is an app thanks to mobile, everything has changed

now, names, concepts, the way we think, how we deal with life's issues, in short everything.

Back then, writing a single line of code was a real pain in the ass, there were no resources for what we were doing, I mean actual resources, not just manuals or short tutorials books, we had to figure out everything ourselves, fixing a bug was a nightmare, sometimes we literally spend weeks trying to fix an error in our code, now programmers use Google to fix it for them! How? You ask, they just type what they're facing on the search engine, and voila... a million answers will pop up in their faces, I wish you were then with me when I had to read 1000 pages of books to finally find nothing, interesting right? No... Books then didn't have search engines, so imagine reading every website

you see online, word by word, this is how it was!

Being a solo programmer did make things more difficult for me, I mean I wasn't alone, I had my younger brother Obada and cousin Derar who also were programmers, but we didn't have this large company that handles everything for us, we had to do everything ourselves, starting by typing the code and ending by struggling with the customer to pay us pennies, yes pennies! No one was taking the software seriously, and more, no one believed that they should pay for a set of code lines on CDs, and if you are of the new generation, CD refers to a compact disk which was a rounded glowing disc we used to store data, just Google it, and be thankful you didn't hear me say floppy disk.

Back then I had to draw the graphics for the software myself, and later prepare the promotional materials like brochures and advertising videos, so above teaching myself how to write the code I had to teach myself how to be a digital graphic designer, a video editor, and everything in between, a million things to do and learn, and to do all that we needed money, we three were committed to start our path independently, we'll take no help from anybody, we said, we needed much though, but we were young and full of hope, long story short, we had to work to provide for our needs, so we worked everywhere doing everything, I recall that Derar ran a small computer repair shop, Obada used to assemble computers on demand, and I worked as a teacher! I even sometimes worked on farms picking up fruits, it

didn't occur to me then to write a "Fruit Ninja" game, damn... if only I could read the future then, I probably would be invested in bitcoin instead! Although it wasn't invented yet, I bet there were better chances than I was so blind to see.

At that stage of my life, I had to take a very serious decision, either I continue studying law and become a lawyer, or to quit and be what I've always dreamed to be, the next Bill Gates, neither eventually! cause to be Gates - I mean a successful programmer who would become a millionaire - I had to quit law school and be a programmer first, but being a programmer wasn't good enough to feed my family, although - at least - I consider myself a programmer and later a digital artist and more,... after all, we are human, we buy

our food, and to do the buying we need the colored note that somehow represents the gold we own, we call it money! Ironically, it now represents ones and zeroes that we own! So neither a lawyer nor Bill Gates and I had to feed my family, I had to take a second major decision in my life, I decided to work abroad, in our culture that's mean to travel to another country, work for some years, collect your salaries and go back to your home country, using the difference in the two countries' currencies rate you'll be able to start a proper project then at your home country, a lot of my friends did that, and maybe thousands of people were doing it for a long time, so why not me too, said the stupid wise man in my head, of course, to most of us in Syria, we usually choose to go to one of the Gulf countries, I've chosen Saudi Arabia for a reason I don't even

want to remember, but any way to Riyadh my plane takeoff.

IN RIYADH

In Riyadh, the capital of Saudi Arabia, I was supposed to start, develop, and run a website, or at least that was the plan, oh... web development was another profession that I taught myself and

mastered it, by the year 2000 everything was about to change, one of the major changes was the web, and to be in any type of business you had to have a website, moreover, it was something to bragging about, everyone wants to own a website, even worse, everyone wants to own a forum for no reason or content, it was that era fad.

Before I went to Riyadh I prepared everything about this project, all the sources I needed, the database, and the graphics even drew on paper what became pages later, I used to draw on paper how my software to be, then after I used digital graphics and still... so I came ready to start, but on the other hand the project owner hadn't had a clue of what he wanted, so we stuck in his ignorance, each time we sat to discuss the project

he had another idea and another scenario and of course, I had to throw everything and start all over again, a delay came after delay, meanwhile he suggested that I work as a digital artist for an advertising agency while he makes up his mind.

Working as a digital artist wasn't a problem for me, I've always loved to draw, I still remember my first painting when I was 3 years old as if it was yesterday, my mom showcases it at one of the school galleries and I still remember how visitors stood there to complement the little boy's art, so in general, it was OK to me to keep on drawing and generating arts, but this is not what I came here to do! Besides, I've always felt that I'm out of time and must hurry up, time was always against me.

I was so frustrated, I thought of going back so many times, and on other hand, I must do something.

At that time, I met the person who later would become my partner for a long time, Fadi was this guy that you had to like from the first time, and I can say that I'm the one you'll hate from first sight, you may like me or not later, but I assure you'll hate me at first! And that was strange, because for a guy like Fadi who is full of energy and passion to hang out with me then, I... who was full of sadness!

Fadi owned a bungee trampoline in an amusement park, and whenever he saw me depressed he would take me there for some fresh air, he used to say, being there for me was the only

reason that kept me on the track that I had to do something, we shared our thoughts about everything and somehow we completed ourselves that way, what I was missing he figured out and vice versa, we are not alike each other in so many ways but that for some reason created this relationship between us!

You see, for all my life I considered Derar to be my only friend, and I say friend because there is no other word that can describe how we feel about each other, yes we are cousins but we never consider ourselves bound because of this, I guess we had to create a new word to describe our relation, we hanged out since we could remember, we spent every weekend at one of our parent's house, we are so connected for we get sick at the

same time with same illness although we live thousands of mile apart now!, in short, we are somehow mirrored, we are the reflection of each other, on the other hand, Fadi wasn't all alike me but in somehow we complete the missing parts in the both of us.

Fadi introduced me to all of his friends, most of them still my friend up today, this was one of his ways to cheer me up, we used to hang out at Khalil's house for hours joining PalTalk rooms, Khalil was and still is the smartest money maker I've ever known, while we spent hours wasting our time on PalTalk Khalil always came up with strange ideas of selling things to people, I said that Khalil can make money selling air in bags, Khalil himself invited me so many time to the club he

worked at to swim and chill up in the steam room and jacuzzi.

In our private time, I and Fadi always discussed our crazy idea about the business, small ideas that were, in fact, the pieces of our big business puzzle later, and till now, I bet if you hear one of them you would probably feel pity for us or call us crazy, our ideas were as crazy as no one will ever think about it, that's because no one thought about merging at least two ideas of them, we were dreamers but also believers.

Those were very hard days with a lot of savors, I've always asked myself, is there a worse thing to come, eventually yes, and as soon as I couldn't imagine, since then I realized that my life

is a bunch of fast-rising ups with faster falling
downs.

BRIGHTNESS COLORS

Forgetting about the project that I was supposed to do in the first place, Fadi and I started our own company later, we called it Brightness colors, it was a media/marketing company, and

after only one month of preparing, we signed our first project contract, and it was a very big project, so big so that no one dares to repeat it till date! Both of us were crazy enough to do it, all you need is courage, said Fadi, we worked day and night to finish that first project, as we were so eager to start our second project, which was even bigger than the first one, we considered the first project just a beginning! Yes, that was the level of our madness, we did not think that anything was impossible or big enough to suit our ambitions.

We were almost a steps far from completing the first project when the second Gulf War took place, and without going into details, Fadi and I were bankrupted at a glance, I've told you, time was always against me!

Fadi and I didn't give up, we worked on smaller projects, but the whole world was struggling I say, the general atmosphere was not conducive to any type of work, in fact, no one wanted to work because everyone considered it an adventure in light of the ongoing situation.

For me losing the company was the biggest shock in my life then, in numbers... actually you can't bear to hear it, every one of our friends was waiting to hear that Fadi and I went crazy about our loss, but somehow, we didn't!

It was time for me to go back home, although I've lost all my money, I still have my life, that I don't want to lose, I know I'm going back empty-handed,

but at least I've tried, so home I was going to start all again.

BACK TO SYRIA

It was a mixed feeling getting back home, I'm happy that I'm finally regrouped with my family, and sad cause I have lost it all, time, money, and efforts, but guess who was there for me? As

always Derar, Derar introduced me to Mr. Ghassan AlKhouri, a very intelligent businessman who ran a GIS software company, the company needed a web mapping software developer, web development was one of my professions, and mapping is something I've learned in no time, I gradually started to recover, things started to improve significantly also, I felt better more and more.

But what do you know, life is full of surprises, Derar decided to work abroad, and to be frank I felt fear and frustration, first, I did not want to lose him, and on the other hand, I was afraid that my experience would be repeated with him, but what I realized, each one of us has his own experience, so maybe he would be luckier than me?

During the following 6-7 years, things began to settle down for me, Fadi and I went back to work together, we established our new company online, it was a web hosting and development company, later we created our first online magazine, and I was directing commercials and short films, our customers vary from celebrities, business, and more, Derar kept visiting me each year, everything was almost going towards better, and before you know it, the Arab Autumn started, they called it spring, it never was because spring never accompany war, so soon war started everywhere, I don't know any more either I'm so lucky with wars or humanity are so stupid that everyone wants to kill everyone!

I had to leave again seeking my small family's safety, after all, you can't predict war nor can you live in it, all that was left to me was my family and I have to keep it safe.

I and my wife and our two kids went to Thailand, a very beautiful Asian country, everything was OK, except for one thing, my kid's school, they couldn't go to school there, and I couldn't afford private schools, so I had to leave again, we choose Egypt, the country of history, everything their reminds you of over 7000 of a great nation story, a very kind and generous people, and above all my kids went to school, this is was so important to me, and there, in Egypt, the story of Wikipedia begin.

WIKIPEDIA

I'm sure you are wondering, what my life story had to do with Wikipedia, I say everything, as you may notice, my life wasn't easy, I've kept losing everything over and over, except one thing,

my experience, and knowledge, I used to consider it my true fortune, so for me, to ask me to give you my experience is like to ask a millionaire to give you his money, it wasn't my style, I don't say we mustn't share knowledge, I'm saying you must work for it as I did in the past, don't just sit there waiting for someone to give it all to you while you are doing nothing! At least this is how I used to think.

In 2014, one of my jobs was to work on the media coverage of the company I worked for after moving to Egypt, my employer said that we must have a Wikipedia page for our company and asked me if I could write one, although I knew what Wikipedia was, I've never written an article there ever, I didn't know the process, the article

structure, adding references, and almost my first article was a profile more than an appropriate article, so eventually, it was refused and later deleted, and I didn't write an article for a long time after.

Although the whole idea of Wikipedia was against my style, no sharing, for some reason I kept reading articles and learning more about how to style an article from time to time, yet I'm not convinced in the whole concept of Wikipedia, remember I was the one who struggled to get paid for writing code for months, I'm the person who's lost everything over and over, and no one ever thought of giving me a hand, moreover a lot of who I knew were trying to push me down more! So yes, Wikipedia wasn't my style, and sharing with

everyone wasn't me.

In 2017 I had a heart attack, I was a smoker for over 17 years, and I was expecting this to happen but not that soon, after the surgery, I went home to start my recovery period, and everything changed in my life after, my whole body starts to fall down, I was in pain everywhere, during the checkups we discovered a lot of problems regarding my health situation, more than I can or want to mention.

Pain is indescribable, the only one who would feel your pain is you.

For almost 20 years I was stuck to my screen doing almost nothing but typing and moving

the mouse, after all, my body will give up on me one day, my body tried so many times to tell me it was time to stop, but I didn't get the messages, so it cracked all at the same time.

After the surgery I spent most of my time in my bed, I barely could move, I'm still, my only amusement was to keep learning and teaching myself, after all, some habits stick to us till we die.

During the next year, I was sure of two things, first I'm going to die soon as my body is not responding well, the second I'm learning faster now, I'm gathering a lot of information in my head, and the idea of not sharing was gradually backing down, I must search for happens in my last days, I kept telling myself.

While surfing the internet for more knowledge I often came across articles on Wikipedia, Wikipedia became an acceptable idea to me now, I have a few days left in my life and maybe it's time for me to rethink, to me now, I do accept sharing, maybe trying to participate in Wikipedia is one thing I could do now, I kept telling myself! Although I still don't have the courage to participate, until one day.

FIRST WIKIPEDIA ARTICLE

Finally, I created an account, wrote an article in Arabic, post it in English on Wikipedia! so again it was deleted, of course, the first rule is: know where you are, I post it again on Arabic Wikipedia,

a week later it was again deleted, the reasons this time were different, some I do understand and I accept and other I can't understand and don't accept, which were related to religion! Yes, Arabic Wikipedia is restricted to this kind of issue, they claim not, but they can't hide the truth, so I quit.

Not for so long, a day after I decided to try again, I began to participate more often on a daily basis in both Arabic and English Wikipedia, the more I participate the more I learned.

My contributions on Arabic Wikipedia were more likely a translation of articles in English Wikipedia, sometimes it was only fixing grammar or typos, I even created an account, using my real full name, for my son so he can participate, later

he lost his password and never wrote again.

For so many reasons I do not want to mention it - but you can read more about it by visiting the Arabic Wikipedia page on English Wikipedia and reading the criticism section, it's way more! - I've always wanted to stop participating in Arabic Wikipedia, but something wasn't right for me, I was about to leave Wikipedia in all until I found out about Egyptian Wikipedia.

EGYPTIAN WIKIPEDIA

In Arabic Wikipedia, I often saw users using a user-box that says: (This user is against Egyptian Wikipedia), at first I didn't fully understand what was going on, until I began to read more about it,

my inner Sherlock Holmes can't calm down, what I discovered that most – not all - of the users on Arabic Wikipedia were against the Egyptian version, more, they consider it a real danger to the Arabic language! Well, this wasn't only in Arabic Wikipedia discussions, it was on articles on other Arabic websites! Fake propaganda was heeled against the project and it was everywhere!

Moreover, the immorality in their way of speaking was disgusting, the way that they described "Ghaly" - the founder of Egyptian Wikipedia – and other users at Egyptian Wikipedia made me feel ashamed, although it wasn't me who said it! Things may not even be acceptable in mythologies, some of the silly users made their own private investigation about Ghaly! Describing

him the Coptic that is leading a project against Islam, really? I didn't know that Arabic Wikipedia was for Muslims only! Nor was Egyptian Wikipedia for Egyptians only, totally nonsense.

So apparently I knew about Ghaly before I even began to participate in Egyptian Wikipedia, Florence Devouard, the former president of the Wikimedia Foundation, stated that the foundation wanted Wikipedians to participate in their native languages, and this is exactly Ghaly's main idea when he proposed the project on 30 March 2008, he declared so many times that Egyptians speak Egyptian all their lives, they never know about Arabic till they went to school! so it's their right to find information in a language that they are familiar with and understand well, this is totally right, you

see, for non-Arabic speakers, you must realize that Arabic is so complicated language, sometimes even to natives, acutely there are no natives these days, each country has its own language, Arabic is mostly used in media and official statements and paper work only, Ironically Arabic defender claims that Arabic is the language used in schools, forgetting that teachers had to explain it using native languages again, moreover, they always use the term, the language of Islam! Again? Really! and let me be clear here, I'm not against the Arabic language or any other language, all I'm saying that no one has the right to force me to use it instead of any other language, me using other language doesn't mean that I'm against you or your language, it doesn't mean I have an agenda against you nor your religions and believes,

remember I used to participate in Arabic Wikipedia, I even wrote books in Arabic, even more, some users, editors and admins in the Arabic Wikipedia demand other contributors to use more restricted words or expressions, otherwise, they will consider the article weakly constructed, so eventually, the article would look as if it was written by a time traveler of 1000 years ago, for someone, like me, who understands Arabic well I can go with that, but how about others, is knowledge for only who can understand historical Arabic, again I'm not saying that all articles on Arabic Wikipedia are like so, I'm just giving an example here.

I was totally convinced of the idea of Egyptian Wikipedia, after all I live in Egypt now and I can offer something, I somehow know more about

the language.

WRITING MORE

I started participating more in Egyptian Wikipedia, alongside participating in both English and Arabic Wikipedia, I was trying my best to fix some errors, restyling templates and more, after

all, Wikipedia is a website and it's easy for me to understand how it's work and improve it whenever I can.

I needed more, I wanted to write more articles, so I started translating articles from English more often, the process was so slow for me, and I don't have enough time in life, I needed more data, I may forget that I was a programmer and data is something I can manipulate! And the source does exist in Wikidata, so I planned my adding process structure, which I'll explain later, I prepared the data and tools to add more articles quicker and started adding articles to Egyptian Wikipedia, meanwhile, I was cooperating with Ghaly in fixing vandalisms in the Egyptian Wikipedia, and with Raafat, who is an

administrator in the Egyptian Wikipedia, on fixing style and creating more required templates, thankfully I can say that they both taught me a lot.

Raafat nominated me to be one of the admins (sysops), which helped me do my jobs faster, things went so good and fast for the next few months, Al-Dandoon (another sysop), Ghaly, Raafat, and I worked day and night on improving Egyptian Wikipedia, other users thankfully worked their best too, and finally, Egyptian Wikipedia hit the 500K articles milestone in a few months only.

Soon after, the stewards (a higher level of admins) decided to reward us at that stage! What they did was:

De-sysop and lock User:Al-Dandoon,

De-sysop User:HitomiAkane,

De-crat User:Ghaly!

Their claims were that our actions were unacceptable, and confirmed that assuming a good faith rule was followed in this case! They found - by holding some kind of a secret investigation - that Ghaly and I acted carelessly and Al-Dandoon considered a global Long Term Abuser! All of Ghaly's attempts to obtain convincing answers to his legitimate questions about what really happened and why came back without a convincing response, although no one can prove it nor want to mention it, I can surely say: look for Arabic hands involved, one of my own suspicions for this was, although stewards involved in the discussions don't speak Egyptian they somehow could understand our local

discussions! Moreover, the way that some admins in Arabic Wikipedia were acting regarding this issue makes you wonder, as long as they stated that they are against this project, why did they waste their time participating in this issue? Moreover, how did they know about what happened when they don't accept the project or participate in it? I can't prove it, but reading between the lines is so easy.

You can refute my allegations or confirm it by searching more in what really happened, but that won't change anything, and despite the confusion of the steward decisions - some of them said that we can apply for the admin's again through them, and some of them said it frankly that they will not allow this to happen - any attempt to

confirm or deny my claims will not, in fact, change anything. The decision was taken to punish us for our efforts!

I was so frustrated, this is not how you reward us for improving the project, it was a depraved way to revenge on our progress, in my opinion, this is not the way you reward Ghaly, especially him, so I decided to quit participating, after all, I'm barely living and can't handle such pressure, until Ghaly came and post, and I quote: "I hope none of you will stop adding and developing this Wikipedia. It is frustrating to see what happened, but I will aim to continue." those words were just enough for me, although I still can't forget what happened, and then when I made my promise to Raafat, since that day I worked my best

to add more and more articles, to fix more, to create more templates and script and to improve Egyptian Wikipedia as long as I could.

By the date, Egyptian Wikipedia is way over 1 million articles, and who knows, if I to live, I may see it reach more.

HOW I DID DO IT?

HOW TO

Now, the main big question here is how I did it? The answer is easy, three things are needed:

1- Data

2- Programming experience

3- Time and efforts

First collecting the data can be easily achieved using the Wikidata project if you could understand it, Wikipedia definition for Wikidata is: Wikidata is a collaboratively edited multilingual knowledge graph hosted by the Wikimedia Foundation. It is a common source of open data that Wikimedia projects such as Wikipedia, and anyone else, can use under the CC0 public domain license. That's simply meant to me a data source.

Most Wikipedia projects uses Wikidata to store a briefs of the articles in form of items, each item consist of properties and values, so, you can see my goal here is to use this data as the base source of my upcoming articles, taking advantage

of it already exists in English Wikipedia, which gives me at minimum millions of potential articles, let's take Bill Gates as an example: Bill Gates item in Wikidata holds the id (Q5284), this item contains numerous properties such as the label (usually means the name and in multi-language), instance of, image, gender, citizenship, occupation and more... each property got its own id which will be used later in the query, sometimes an item holds less properties sometimes more than other item, some items uses the same properties, so understanding Wikidata is the first step, understanding how each item is represented using properties and how properties work, this all so important before moving to the next step, next I need to search for a group of items that uses the same properties, let's say I need to search for all

football players that do have citizenship, date of birth, gender, and more, so I can create articles about football players, to do that I need a tool to do the query, which do exist at the URL:

<https://query.wikidata.org/>

Wikidata Query Service is a query engine to perform queries on Wikidata database, it uses SPARQL, an RDF query language -that is, a semantic query language for databases- able to retrieve and manipulate data stored in Resource Description Framework (RDF), and so querying football players using this tool could be done by:

- 1- Knowing the association football player property id as it will be used to group all items.
- 2- Knowing what properties the items should have

and their ids, and keep in mind, the more properties each item have the richer the article would be but the fewer results I'll get.

3- Get the results in English.

4- Limit my query to a suitable number, a large number may cause query time limitation error.

So to achieve that I can run the following SQL in Wikidata Query Service website:

```
SELECT DISTINCT
?item ?itemLabel ?itemsexorgenderLabel ?itemcountryofcitizenshipLabel
?itemplaceofbirthLabel ?itemdateofbirth
WHERE {
?item wdt:P106 wd:Q937857. # occupation: association football player
OPTIONAL {?item wdt:P21 ?itemsexorgender.}
OPTIONAL {?item wdt:P27 ?itemcountryofcitizenship.}
OPTIONAL {?item wdt:P19 ?itemplaceofbirth.}
OPTIONAL {?item wdt:P569 ?itemdateofbirth.}
SERVICE wikibase:label { bd:serviceParam wikibase:language
"en,[AUTO_LANGUAGE]".}
```

```
}
```

```
LIMIT 100
```

Running the above SQL will hopefully return a set of 100 football player with the minimum required data for creating an article which are: Wikidata item ID, name, gender, citizenship, place of birth, and date of birth, of course, I can add other properties to my query, the result of the query then could be exported to many formats, JSON is one of them and my favorite one.

Two other things are important, first, I need to make sure that each item does have an article on English Wikipedia not just any random items on Wikidata, and the second thing is to make sure that I didn't already have this item added to my target Wikipedia so I'll end up having duplicated articles,

to do so I can use the "FILTER EXISTS" and "FILTER NOT EXISTS" enclosures in my previous query, so I can edit it to be like follows:

```
SELECT DISTINCT
?item ?itemLabel ?itemsexorgenderLabel ?itemcountryofcitizenshipLabel
?itemplaceofbirthLabel ?itemdateofbirth
WHERE {
?item wdt:P106 wd:Q937857. # occupation: association football player
OPTIONAL {?item wdt:P21 ?itemsexorgender.}
OPTIONAL {?item wdt:P27 ?itemcountryofcitizenship.}
OPTIONAL {?item wdt:P19 ?itemplaceofbirth.}
OPTIONAL {?item wdt:P569 ?itemdateofbirth.}
SERVICE wikibase:label { bd:serviceParam wikibase:language
"en,[AUTO_LANGUAGE]".}
FILTER NOT EXISTS {
?article schema:about ?item .
?article schema:isPartOf <https://XXX.wikipedia.org/> .
}
FILTER EXISTS {
?article schema:about ?item .
?article schema:isPartOf <https://en.wikipedia.org/> .
}
}
LIMIT 100
```

Notice that in the "FILTER EXIST" enclosure I defined the English Wikipedia, so it will make sure that the results only contain items that already exist in English Wikipedia, and in the "FILTER NOT EXIST" enclosure I should define my target Wikipedia, by replacing the "XXX" prefix according to my target Wikipedia prefix, for example, fr for French Wikipedia.

That's of course depends on, that my target Wikipedia does use Wikidata and connect newly created articles to their appropriate Wikidata items! Otherwise, this process will only make sure to bring results that exist in English Wikipedia regardless of their existence in my target Wikipedia.

Building the correct query is the key to obtaining useful data for building a rich article, although SPARQL may look difficult, in fact, it's easy to understand as long those who use it know what SQL is, and in addition, there are a set of useful samples and tutorials at:

https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Wikidata_Query_Help

There are even tools for visually building queries like:

<https://wd-query-builder.toolforge.org/>

Now I've got my pure data, I need to format

it in article form, so, for each line of the result I do replacement and create the article(s) core content, for example:

[label] (**[date of birth]**), **[gender]** is a football player from **[citizenship]**, **[gender]** was born at **[date of birth]** in **[place of birth]**.

It's obvious from the above line that it's almost like filling a template using the data extracted from each line, some minor tweaks are still required such as converting the gender to he/she and so on, of course, I still need to add suitable infoboxes, categories, and references and do the translation to my Wikipedia desired language.

This looks like a lot of work, and there is

where I needed my programming skill.

After exporting the result to the preferred format, I can use a programming language to loop each line of the result and do the replacement, do the translation, and add the extra information such as the categories, then print out the final result which is the article in my case, usually, I do this using PHP, so first I load the JSON file and convert it to an array like this example:

```
$dataFile = file_get_contents("path to json file");  
$jsonArray = json_decode($dataFile, true);
```

Then I go through each array item and create the content, something like:

```
foreach ($jsonArray as $key => $value) {
```

How I wrote a million Wikipedia articles

```
echo "{{infobox}}\n";
echo "" . $jsonArray[$key]['incidentLabel'] . "" ("";
echo $jsonArray[$key]['dateOfBirth'] . ") is a football player from ";
echo $jsonArray[$key]['citizenshipLabel'] . ". ";
echo ...
..
.
}
```

Also using PHP I add missing extra info such as infobox and categories, do the needed translations and so on.

All I have to do next is: copy the article content, go to Wikipedia, create a new article, past it there and publish it, connect the Wikidata item with the newly created article, and this is the time and efforts part of the process.

This process could be even boosted up

using MediaWiki API to add articles on the fly while creating its content, the MediaWiki Action API is a web service that allows access to some wiki-features like authentication, page operations, and search, in my case, and I mainly focus on creating/editing pages API.

More information on that at this URL:

<https://www.mediawiki.org/wiki/API:Edit>

WHAT TO WRITE

In my giving SQL example I assumed searching for football players by grouping items by this profession, but what if I need to write about another profession? Then all I need to do is to replace the

ID of this profession (which is Q937857 in the example) with desired profession ID, and the main question here is where I can find professions IDs?

Simply again, using the same tool I can query professions, and to save my time, I made the query, which is as follows:

```
SELECT DISTINCT
?item ?itemLabel
WHERE {
  ?item wdt:P31 wd:Q28640. # Instance of occupation
  SERVICE wikibase:label { bd:serviceParam wikibase:language
"en,[AUTO_LANGUAGE]".}
  FILTER EXISTS {
    ?article schema:about ?item .
    ?article schema:isPartOf <https://en.wikipedia.org/> .
  }
}
LIMIT 1000
```

This query will result in a 1000 occupation with their associated Wikidata IDs, simply chose an ID, replace it with the football player occupation ID in the first SQL, add other relive properties as needed and run the query.

So let's say I've chosen to go with a computer scientist profession which holds the ID Q82594, the SQL will look something like this:

```
SELECT DISTINCT
?item ?itemLabel ?itemsexorgenderLabel ?itemcountryofcitizenshipLabel
?itemplaceofbirthLabel ?itemdateofbirth
WHERE {
?item wdt:P106 wd:Q82594. # occupation: computer scientist
OPTIONAL {?item wdt:P21 ?itemsexorgender.}
OPTIONAL {?item wdt:P27 ?itemcountryofcitizenship.}
OPTIONAL {?item wdt:P19 ?itemplaceofbirth.}
OPTIONAL {?item wdt:P569 ?itemdateofbirth.}
SERVICE wikibase:label { bd:serviceParam wikibase:language
"en,[AUTO_LANGUAGE]". }
```

```
FILTER NOT EXISTS {  
  ?article schema:about ?item .  
  ?article schema:isPartOf <https://XXX.wikipedia.org/> .  
}  
FILTER EXISTS {  
  ?article schema:about ?item .  
  ?article schema:isPartOf <https://en.wikipedia.org/> .  
}  
}  
LIMIT 100
```

Of course, not all article should be biographical, what if you want to write about something else human professions? Then I can replace the "occupation" term with "instance of" term, which tells the query engine to group all items that have the same instance, of course the associated properties should also be change, because for an example, an airport item can't have the "date of birth" property, so speaking

about airports, a simple SQL could be something like:

```
SELECT DISTINCT
?item ?itemLabel ?countryLabel ?locatedInadministrativeLabel ?inceptionDate
?seaLevelElevation ?IATALabel ?ICAOLabel
WHERE {
?item wdt:P31 wd:Q1248784; #instance of airport
OPTIONAL {?item wdt:P17 ?country.}
OPTIONAL {?item wdt:P131 ?locatedInadministrative.}
OPTIONAL {?item wdt:P571 ?inceptionDate.}
OPTIONAL {?item wdt:P2044 ?seaLevelElevation.}
OPTIONAL {?item wdt:P238 ?IATA.}
OPTIONAL {?item wdt:P239 ?ICAO.}
SERVICE wikibase:label { bd:serviceParam wikibase:language
"en,[AUTO_LANGUAGE]".}
FILTER NOT EXISTS {
?article schema:about ?item .
?article schema:isPartOf <https://XXX.wikipedia.org/> .
}
FILTER EXISTS {
?article schema:about ?item .
?article schema:isPartOf <https://en.wikipedia.org/> .
}
}
```

limit 1000

This SQL will result in 1000 airports with available properties.

So, the point in focus is that I must deal with each item according to its associated properties.

Fortunately, Wikidata Query Service website contains plenty of useful examples that explain a lot about the SQL syntaxes and how to achieve specific queries, also the Wikidata Query Help website offers a lot of useful information about almost everything.

FINALLY

FINAL WORDS

As you can see there is no magical solution here, the process is time-consuming, and above that, it needs a strong determination, telling you my story was my way of telling you to keep trying and

never give up, nothing comes so easy, also nearly nothing is impossible, although I don't know how many days I have left, I still do what I love, no matter what you face in your life, just keep trying, don't let anybody makes you feel downed, not everyone will be happy to see your success, all you have to do is to keep doing what you love and enjoy doing it.

I may be was aggressively talking about what happened in Egyptian Wikipedia by the stewards, after all this my point of view and only represents me, and it may be not even acceptable by my colleagues in Egyptian Wikipedia, all I'm saying that we worked as volunteers, corrupting Wikipedia wasn't our goal, never, it's a project we all are trying to improve and what happened was

unacceptable to me, if the stewards thought that something was incorrect they could simply inform us before taking such an action, I say, but anyhow, although I can't prove anything, I just want to set the record straight about what happened for the future, we do all make mistakes, we all do miss judge things, nothing in life is perfect nor ideal.

This is not to let you down, or a call to withhold you from participating in any kind of projects that could help anybody, even in a thousand years yet to come, for me, it was a great experience to participate in Wikipedia, achieving more than a million article by the date of writing this book is something I wasn't expecting to happen - so fast - but I've always wanted since the beginning, and I did it, I may not be able to

participate anymore, either way, it was nice being there, see you all on the next side.

Maher Asaad Baker

AKA **HitomiAkane** (Wikipedia user name)

**"If corruption is a disease
transparency is essential part
of its treatment"**

Kofi Annan

FURTHER READING

You may be interested in taking a look at those links, which may help explain my point of view:

- 1- https://arz.wikipedia.org/w/index.php?oldid=4663693#Steward_removal_of_flags_on_arzwiki
- 2- <https://meta.wikimedia.org/w/index.php?oldid=20130801#Miscellaneous>

How I wrote a million Wikipedia articles

s_requests

- 3- https://arz.wikipedia.org/w/index.php?oldid=4663693#Hope_for_the_future
- 4- <https://en.wikipedia.org/w/index.php?oldid=1007733242>
- 5- <https://en.wikinews.org/w/index.php?oldid=1195289>

DISCLAIMER

This is a work of nonfiction. No names have been changed, no characters invented, no events fabricated.

This work represents the author's point of view only.

Unless otherwise indicated, all the names, characters, businesses, places, events, and incidents in this book are used for reference only and they are fully respected, no harm or disrespect was ever meant.

Although the publisher and the author have made every effort to ensure that the information in this book was correct at press time and while this publication is designed to provide accurate information in regard to the subject matter covered, the publisher and the author assume no responsibility for errors, inaccuracies, omissions, or any other inconsistencies herein and hereby

disclaim any liability to any party for any loss, damage, or disruption caused by errors or omissions, whether such errors or omissions result from negligence, accident, or any other cause.

No animals were harmed during the making of this work.

ABOUT THE AUTHOR

Maher Asaad Baker (In Arabic: **ماهر أسعد بكر**) is a Syrian musician, VFX & Graphic artist, director, and author, he was born in Damascus in 1977. Since his teens, he has been building up his

career, starting by developing applications and websites while exploring various types of media-creating paths.

He started his career in 1997 with a dream of being one of the most well-known artists in the world.

As music plays a very important role in bringing people together, it also makes them live life with all their emotions exposed. He makes mixed emotional music so everyone can express their feelings, he is known for his music being produced, mixed, mastered, engineered, written, and performed himself. Most of his music contains hip hop and R&B elements.

Reading was always a part of his life as he was always surrounded by his father's books, but his writing ability didn't develop until a later age as his most time was occupied with other things such as developing, writing songs and music, or in media projects production, he is most known for his book "How I wrote a million Wikipedia articles" and a novel entitled "Becoming the man".

